

Appendix C

Statistical Considerations

The monthly volumes and prices of natural gas to residential, commercial, and industrial consumers presented in this report by State are estimated from data reported on the Form EIA-857, "Monthly Report of Natural Gas Purchases and Deliveries to Consumers."

Monthly prices in select States are supplemented with data from the Form EIA-910, "Monthly Natural Gas Marketer Survey." (See Appendix B, "Data Sources," for a description of these Forms.)

A description of the sample design and the estimation procedures is given below.

Sample Design

The Form EIA-857 implements a monthly sample survey of companies delivering natural gas to consumers. It includes inter- and intrastate pipeline companies and local distribution companies.

The survey provides data that are used each month to estimate the volume of natural gas delivered and the price for onsystem sales of natural gas by State to three consumer sectors—residential, commercial, and industrial.

Monthly deliveries and prices of natural gas to the electric power sector through 2007 were reported on the Form EIA-906, "Power Plant Report"; Form EIA-920, "Combined Heat and Power Plant Report"; Form FERC-423, "Monthly Report of Costs and Quality of Fuels for Electric Plants"; and Form EIA-423, "Monthly Cost and Quality of Fuels for Electric Plant Report."

Beginning with 2008 data, the Form EIA-923, "Power Plant Operations Report," replaced these sources.

Sample Universe. The sample in use for 2011 was selected from a universe of 1,539 companies. These companies were respondents to the Form EIA-176, "Annual Report of Natural and Supplemental Gas Supply and Disposition," for reporting year 2009, who reported sales or deliveries to consumers in the residential, commercial, or industrial sectors. (See Appendix B for a description of the Form EIA-176.)

Sampling Plan. The goal is a sample that provides estimates of monthly natural gas consumption by the three consuming sectors within each State and the District of Columbia. A stratified sample using a single stage design and systematic selection with probability proportional to size is implemented to achieve this.

The measure of size is the total volume of natural gas delivered in a State to the residential, commercial, and industrial sectors by the company in 2009. There are two strata—companies selected with certainty and companies selected systematically under probability proportional to size.

When the design described above was first constructed, calculations suggested a sample consisting of approximately 25 percent of all companies yields reasonably precise estimates. The sample for 2011 was selected individually in each State and the District of Columbia. It consists of 401 companies.

Certainty Stratum. Since estimates are needed for each of the 50 States and the District of Columbia, the strata are established independently within each State. In 16 States and the District of Columbia where sampling is not feasible due to small numbers of companies and/or small volumes of gas deliveries, all companies are selected.

The 16 States with all companies selected are: Alaska, Connecticut, Delaware, Hawaii, Idaho, Maine, New Hampshire, New Jersey, Nevada, North Dakota, Oregon, Rhode Island, South Dakota, Utah, Vermont, and Washington.

For the remaining States, the volumes of (1) industrial sales and deliveries, (2) combined commercial and residential sales and deliveries, and (3) the sum of all three sector sales and deliveries are computed. Companies with natural gas deliveries to (1), (2), or (3) above a certain cutoff level are selected with certainty.

Since a few large companies often account for most of the natural gas delivered within a State, this ensures those companies' inclusion in the sample. The formula for determining certainty is applied independently for (1), (2), or (3).

The companies exceeding the cutoffs for (1), (2), or (3) form the initial selections for the certainty stratum. The rest of the certainty stratum consists of companies in jurisdictions where sampling is not feasible or of companies that are force selected to allow for the drawing of a more representative sample in the noncertainty stratum.

All companies with natural gas deliveries in sector j greater than the cut-off value ($C_{.j}$) are included in the certainty stratum. The formula for $C_{.j}$ is:

$$C_{.j} = \frac{X_{.j}}{2n} \quad (1)$$

where:

$C_{.j}$ = cutoff value for consumer sector j ,

n = target sample size to be selected for the State, 25 percent of the companies in the State,

$X_{.j}$ = the sum within State of annual gas volumes in consumer sector j ,

Noncertainty Stratum. All other companies form the noncertainty stratum. They are systematically selected with probability proportional to size. The measure of size for each company is the total volume of gas sales to all consumer sectors (X_i).

The number of companies to be selected from the noncertainty stratum is calculated for each State. The minimum noncertainty stratum size is two.

The formula for selecting the number of noncertainty stratum companies is:

$$m = n \frac{X_2}{X_{..}} \quad (2)$$

where:

m = the sample size for the noncertainty stratum within a State,

X_2 = the sum within State of the X_i for all companies in the noncertainty stratum.

$X_{..}$ = the sum within State of annual gas volumes in all consumer sectors.

Companies are listed in ascending order according to their measure of size. Then a cumulative measure of size in the stratum is calculated for each company. The cumulative measure of size is the sum of the measures of size for that company and all preceding companies on the list. An interval of width I for selecting the companies systematically was calculated.

A uniform random number R is selected between zero and $\left(I = \frac{X_2}{m}\right)I$. The first sampled company is the first company on the list to have a cumulative measure of size greater than R .

The second company selected was the first company on the list to have a cumulative measure of size greater than $R + I$. $R + I$ is increased by I to determine the third company to be selected. This procedure is repeated until the entire sample is drawn.

Substrata. In four States, the noncertainty stratum is divided into substrata to ensure that gas in each consumer sector can be estimated. The systematic sample with probability proportional to size design described above is applied independently in each substratum.

The methods for determining the substratum sample size and calculating the substratum interval for sample selection are the same as described above for the noncertainty stratum, except that X_2 is the sum within State of the X_i for only those companies in the substratum.

These substrata are defined only for the purpose of sample selection. They are:

Kansas, Louisiana, Oklahoma, and Texas. The substrata are (1) companies delivering gas only to industrial consumers and (2) all other companies.

Estimation Procedures

Estimates of Volumes. To estimate the total gas sales and deliveries for the State, a ratio estimator is applied to the reported volumes in each State for the sampled companies.

Ratio estimators are calculated for each consumer sector – residential, commercial, and industrial – in each State where companies are sampled using annual data from the most recent submission of Form EIA-176.

The formula for calculating the ratio estimator ($T_{.j}$) for the volume of gas in a State for consumer sector j is:

$$T_{.j} = \frac{\sum_{i \in S} y_{ij}}{\sum_{i \in S} x_{ij}} \quad (3)$$

where:

y_{ij} = the monthly volume within a State for sector j and company i that is an element of the EIA-857 sample, s ,

x_{ij} = the annual volume within a State for sector j and company i that is an element of the EIA-857 sample, s .

The ratio estimator is applied as follows:

$$V_{.j} = T_{.j} * X_{.j} \quad (4)$$

where:

$V_{.j}$ = the estimate of monthly gas volumes in a State for consumer sector j ,

$X_{.j}$ = the annual volume within a State for sector j for all companies.

For most respondents, the estimation of residential and commercial consumption requires one additional step than the industrial sector. In this extra step the company level reported volumes are adjusted using sendout. Sendout, which was added to Form EIA-857 for the August 2010 report month, is the total volume of gas dispatched for delivery during a calendar month. It should be noted that the adjusted volumes for the residential and commercial sector are used in equation 4.

The adjustment is implemented by taking reported sendout and removing gas that is consumed in company operations and distribution use as well as gas delivered to electric power, industrial, and vehicle use sectors. The remaining gas is allocated to the residential and commercial sectors proportionally based on reported totals for the report month. For a small number of companies, the adjustment to sendout is applied to all sectors in order to better represent their billing practices. In these cases, sendout is applied to all end-use sectors proportionally instead of just to residential and commercial.

The formulas for adjusting the reported residential and commercial volumes at the company level are:

$$Radj = (SO * (1 - Oth) - EP - IN) * \frac{R}{R + C} \quad (5)$$

$$Cadj = (SO * (1 - Oth) - EP - IN) * \frac{C}{R + C} \quad (6)$$

where:

R_{adj} = sendout adjusted residential volume,

C_{adj} = sendout adjusted commercial volume,

SO = sendout volume,

Oth = ratio of annual distribution use and vehicle fuel to total deliveries,

EP = reported electric power volume,

IN = reported industrial volume,

R = reported residential volume,

C = reported commercial volume.

Computation of Natural Gas Prices. The natural gas volumes that are included in the computation of prices represent only those volumes associated with natural gas sales by natural gas companies except as explained below.

The price of natural gas within a State for sector j is calculated as follows:

$$P_j = \frac{R_j}{V_j} \quad (7)$$

where:

P_j = the average price for gas sales within a State for consumer sector j ,

R_j = the reported revenue from natural gas sales within a State for consumer sector j ,

V_j = the reported volume of natural gas sales within a State for consumer sector j .

All average prices are weighted by their corresponding sales volume estimates when national average prices are computed.

The monthly average prices of natural gas to residential and commercial consumers in Georgia, Maryland, New York, Ohio, and Pennsylvania are monthly average prices of natural gas, based on total sales (sales by local distribution companies and natural gas marketers).

Beginning in January 2005, the EIA-910 added the States of Florida, Illinois, Massachusetts, Michigan, New Jersey, Virginia, West Virginia, and the District of Columbia. Residential and commercial prices represent total deliveries of gas sold to customers in those States as the quality of data collected on the EIA-910 improved with time.

Volumes of gas delivered for the account of others to these consumer sectors are not included in the State or national average prices except in these States.

The price of natural gas in the residential and commercial sectors where EIA-910 data are used is calculated as follows:

$$P_c = \left[\left(\frac{R_s}{V_s} \right) * \left(\frac{V_s}{V_s + V_t} \right) \right] + \left[\left(\frac{Rm_s}{Vm_s} \right) * \left(\frac{V_t}{V_s + V_t} \right) \right] \quad (8)$$

where:

P_c = the combined average price for gas sales by local distribution companies and marketers within the State in sector s (residential or commercial),

R_s = the reported revenue from natural gas sales by local distribution companies within the State in sector s (residential or commercial),

V_s = the reported volume of natural gas sales by local distribution companies within the State in sector s (residential or commercial),

V_t = the reported volume of natural gas transported by local distribution companies for marketers within the State in sector s (residential or commercial),

Rm_s = the reported revenue from natural gas sales by marketers within the State in sector s (residential or commercial),

Vm_s = the reported volume of natural gas sales by a marketer within the State in sector s (residential or commercial).

Table 22 shows the percent of the total State volume that represents volumes from natural gas sales to the residential, commercial, and industrial sectors. This table may be helpful in evaluating residential, commercial, and industrial price data.

Estimation for Nonrespondents and Edit Failures. A volume for each delivered and transported consumer category is imputed for companies that fail to respond in time for inclusion in the published estimates (unit nonresponse) or for which reported volumes have failed the edit and not been confirmed or corrected (item nonresponse).

In the case of unit nonresponse, the imputed volumes for the residential, commercial, and industrial sectors are derived through a multi-stage procedure:

(1) Prediction of monthly volumes for the commercial, industrial, and residential sectors within Census Division. Census Division refers to the nine divisions into which the U.S. Bureau of the Census groups the 50 States and the District of Columbia for reporting and analysis purposes.

For the commercial and residential sectors, the predicted division volume for a month depends on the heating degree days reported by the National Oceanic and Atmospheric Administration (NOAA) for that month within the Census Division as well as a monthly variable that captures historical trends particular to that month.

For the industrial sector, the predicted division volume for a month depends on the same month from the previous year as well as year-to-year differences from the previous months.

The coefficients are estimated via ordinary least squares multiple linear regression. The source is a database of monthly sector volumes for the 5 years ending December 31 of the immediately prior calendar year. Coefficient estimation is restricted to companies reporting continuously during the 5 years.

(2) Allocating the monthly sector volume for a particular respondent based on the respondent's share of that sector volume in the latest Form EIA-176 survey.

Once the predicted division volume for a sector is obtained, it is multiplied by an allocation factor to obtain the imputed sector volume for a respondent.

The allocation factor is the ratio of that respondent's sector volume to the total of all such sector volumes as reported in the latest Form EIA-176 survey.

(3) The final piece of data to be imputed is sendout. As reported sendout includes volumes consumed distributing natural gas as well as deliveries to vehicle fuel, residential, commercial, industrial, and electric power sectors, the imputation method aggregates estimates for each of these components. Distribution use and vehicle fuel are only collected annually, thus monthly totals are estimated based on annual ratios against total deliveries. Industrial and electric power sectors are both calculated using the industrial method that was previously described. The estimation of the residential and commercial components uses a company specific regression model with localized heating degree data.

In the case of item nonresponse for sendout, the residential and commercial sectors default to the reported values. Item nonresponse for all other items is handled in the same manner as unit nonresponse.

Estimation of Revenue. The company's previous month's sector-specific price is multiplied by the corresponding sales volume to impute revenue for that sector.

Final Revisions

Adjusting Monthly Data to Annual Data. After the annual data reported on the Form EIA-176 have been submitted, edited, and prepared for publication in the *Natural Gas Annual*, revisions are made to monthly data.

The revisions are made to the volumes and prices of natural gas delivered to consumers that have appeared in the *Natural Gas Monthly* (NGM) to match them to the annual values appearing in the *Natural Gas Annual*.

The revised monthly estimates allocate the difference between the sum of monthly estimates and the annual reports, according to the distribution of the estimated values across the months.

Before the final revisions are made, changes or additions to submitted data received after publication of the monthly estimate and not sufficiently large to require a revision to be published in the NGM, are used to derive an updated estimate of monthly consumption and revenues for each State's residential, commercial, or industrial natural gas consumption.

For each State, two numbers are revised, the estimated consumption and the estimated price per thousand cubic feet.

The formula for revising the estimated consumption is:

$$V_{jm}^* = V_{jm} \left(\frac{V_{ja}}{V_{jm}'} \right) \quad (9)$$

where:

V_{jm}^* = the final volume estimate for month m in consumer sector j,

V_{jm} = the estimated volume for month m in consumer sector j,

V_{ja} = the volume for the year reported on Form EIA-176,

V_{jm}' = the annual sum of estimated monthly volumes.

The price is calculated as described above in the Estimation Procedures section, using the final revised consumption estimate and a revised revenue estimate.

The formula for revising the estimated revenue is:

$$R_{jm}^* = R_{jm} \left(\frac{R_{ja}}{R_{jm}'} \right) \quad (10)$$

where:

R_{jm}^* = the final revenue estimate for month m in consumer sector j,

R_{jm} = the estimated revenue for month m in consumer sector j,

R_{ja} = the revenue for the year reported on Form EIA-176,

R_{jm}' = the annual sum of estimated monthly revenues.

Revision of Volumes and Prices for Deliveries to Electric Power Sector. Revisions to monthly deliveries to the electric power sector are published throughout the year as they become available.

Reliability of Monthly Data

The monthly data published in this report are subject to two sources of error -- nonsampling error and sampling error. Nonsampling errors occur in the collection and processing of the data.

See the discussion of the Form EIA-857 in Appendix B for a description of nonsampling errors for monthly data.

Sampling error may be defined as the difference between the results obtained from a sample and the results that a complete enumeration would provide. The standard error statistic is a measurement of sampling error.

Standard Errors. A standard error of an estimate is a statistical measure that indicates how the estimate from the sample compares to the result from a complete enumeration.

The standard errors for monthly natural gas volume and price estimates by State are given in Table C1. They are constructed so that over the long run the true value is within two standard errors of its estimate 95% of the time. For volume estimates, the relative standard error (RSE) is published as a percent. Standard errors of prices are expressed in dollars.

The relative standard error of a natural gas volume estimate is the square root of the variance of the estimate divided by the published volume. The

formula for calculating the variance of the volume estimate is:

$$Var(V_j) = \sigma_\epsilon^2 \sum_{i=n+1}^N x_{ij} + \left(\sigma_\epsilon^2 / \sum_{i=1}^n x_{ij} \right) \left(\sum_{i=n+1}^N x_{ij} \right)^2 \quad (11)$$

where:

$V_{.j}$ = the estimate of monthly gas volumes in a State for sector j,

$\sigma_e^2 = \sum_{i=1}^n \sigma_{0i}^2 / (n - 1)$, the sum of squared

residuals,

$\sigma_{0i} = \frac{(y_{ij} - T_{.j} x_{ij})}{\sqrt{x_{ij}}}$, the weighted residual for company i,

$T_{.j}$ = the ratio estimator defined in equation 3,

x_{ij} = the reported annual volume for company i in sector j,

y_{ij} = the reported monthly volume for company i in sector j,

n = sample size ($i > n$ implies company i not in sample),

N = frame size.

As the price of natural gas is calculated as a quotient of revenue to volume, the standard error of natural gas prices is derived using the variance properties of a quotient. The variance of revenue is calculated in the same manner as the volumes described in equation 11, with revenue substituted for volume. The formula is for calculating the standard error of prices is below:

$$Var(P_i) = Var\left(\frac{R_i}{V_i}\right) = \left(\frac{E(R_i)}{E(V_i)}\right)^2 * \left(\frac{Var(R_i)}{E(R_i)^2} + \frac{Var(V_i)}{E(V_i)^2} - 2 \left(\frac{Cov(R_i, V_i)}{E(R_i) * E(V_i)} \right) \right) \quad (12)$$

where:

P_i = natural gas price in sector i,

R_i = natural gas revenue in sector i,

V_i = natural gas volume in sector i,

$E(V_i)$ = expected value of volume in sector i, which is the $T_{.j} * X_{ij}$ product in equation 4,

$E(R_i)$ = expected value of revenue in sector i calculated in the same manner as the $T_{.j} * X_{ij}$ product in equation 4,

$Var(V_i)$ = the variance of volume in sector i calculated using equation 11,

$Var(R_i)$ = the variance of revenue in sector i calculated in the same manner as equation 11,

$Cov(V_i, R_i)$ = the covariance between revenue and volume in sector.

Table C1. Standard Error for Natural Gas Deliveries and Price by Consumers, by State, December 2011

State	Volume Relative Standard Error (Percent)				Price (Dollars per Thousand Cubic Feet)		
	Residential	Commercial	Industrial	Total	Residential	Commercial	Industrial
Alabama	1.2	1.1	1.6	1.0	0.14	0.15	0.15
Alaska	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Arizona	0.2	0.3	0.3	0.1	0.03	0.03	0.02
Arkansas	0.4	0.2	0.1	0.1	0.02	0.03	0.16
California	0.4	0.5	0.2	0.2	0.04	0.05	0.04
Colorado	0.5	0.8	NA	NA	0.06	0.05	NA
Connecticut	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Delaware	0.0	0.0	NA	NA	0.00	0.00	0.00
District of Columbia	0.0	0.0	0.0	0.0	0.00	NA	0.00
Florida	1.2	0.6	2.4	1.3	0.18	0.17	0.72
Georgia	0.2	0.8	1.8	0.7	0.24	0.13	0.20
Hawaii	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Idaho	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Illinois	0.2	0.2	0.7	0.2	0.04	0.06	0.21
Indiana	0.1	0.5	1.4	0.8	0.04	0.10	NA
Iowa	0.6	0.3	1.7	0.8	0.07	0.06	0.29
Kansas	1.9	1.4	1.1	1.0	0.16	0.19	NA
Kentucky	0.3	0.6	0.6	0.3	0.09	0.10	0.04
Louisiana	1.3	0.9	1.6	1.5	0.08	0.15	NA
Maine	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Maryland	0.1	0.4	0.2	0.2	0.01	0.07	0.18
Massachusetts	1.2	1.9	NA	NA	0.16	0.29	0.34
Michigan	0.0	0.3	0.7	0.1	0.02	0.04	0.20
Minnesota	0.5	0.6	2.4	0.8	0.06	0.05	0.20
Mississippi	1.8	2.1	3.9	2.3	0.14	NA	0.05
Missouri	0.1	0.7	1.1	0.3	0.03	0.03	0.31
Montana	0.5	0.2	0.0	0.2	0.01	0.04	0.00
Nebraska	7.0	0.8	NA	NA	0.80	0.11	0.28
Nevada	0.0	0.0	0.0	0.0	0.00	0.00	0.00
New Hampshire	0.0	0.0	0.0	0.0	0.00	0.00	0.00
New Jersey	0.0	0.0	0.0	0.0	NA	0.00	0.00
New Mexico	5.4	5.2	0.0	3.3	0.25	0.23	0.00
New York	0.1	0.3	0.9	0.1	0.05	0.10	0.15
North Carolina	0.1	0.4	0.3	0.1	NA	0.08	0.45
North Dakota	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Ohio	0.4	0.7	2.1	0.7	0.19	0.73	0.65
Oklahoma	0.7	0.4	0.6	0.4	0.02	0.07	0.09
Oregon	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Pennsylvania	0.1	0.1	0.6	0.2	0.03	0.03	0.45
Rhode Island	0.0	0.0	NA	NA	0.00	0.00	0.00
South Carolina	0.5	1.1	1.4	0.8	NA	0.08	0.05
South Dakota	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Tennessee	1.0	1.3	1.7	0.8	0.08	0.14	0.32
Texas	1.5	1.0	0.7	0.6	0.12	0.11	0.06
Utah	NA	NA	0.0	NA	0.00	0.00	0.00
Vermont	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Virginia	2.6	2.5	2.3	1.5	0.31	NA	0.60
Washington	0.0	0.0	0.0	0.0	0.00	0.00	0.00
West Virginia	0.4	0.8	0.1	0.3	0.09	0.08	0.04
Wisconsin	0.2	0.3	0.4	0.2	0.02	0.05	0.14
Wyoming	0.7	2.2	1.3	NA	0.04	0.25	NA
Total	0.1	0.1	0.3	0.1	0.02	0.02	0.05

NA Not available.

Source: Energy Information Administration (EIA): Form EIA-857, "Monthly Report of Natural Gas Purchases and Deliveries to Consumers."